

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001  
TIME: 08:56:25

Input Set : A:\Seqlist.txt  
Output Set: N:\CRF3\12132001\I003608.raw

Does Not Comply  
Corrected Diskette Needed

4 <110> APPLICANT: Bartha, Gabor  
5 Walker, Michael  
W--> 7 *11207 7 move up*  
W--> 8 METHODS FOR ANALYZING GENE EXPRESSION PATTERNS  
12 <130> FILE REFERENCE: ICYTP012  
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/003,608  
C--> 14 <141> CURRENT FILING DATE: 2001-11-01  
14 <150> PRIOR APPLICATION NUMBER: 60/245,081  
15 <151> PRIOR FILING DATE: 2000-11-01  
17 <160> NUMBER OF SEQ ID NOS: 30  
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
21 <210> SEQ ID NO: 1  
22 <211> LENGTH: 4588  
23 <212> TYPE: DNA  
24 <213> ORGANISM: Human  
26 <400> SEQUENCE: 1  
27 gcggcgccca agagagatca cccccccagc cgaccctgcc agcgagcgag cccgacccca 60  
28 ggcggtccatg gagcgctcgcc tccgcccggg ccctgccccg acccccgccct gcggcgcggc 120  
29 tctgccttg accaggactt gggactttgc gaaaggatcg cggggcccg agaggttaacc 180  
30 gccgcgcctc ccggagaggt gttggagagc acaatggctg aacaagtcct tcctcaggct 240  
31 ttgtatttga gcaatatgcy gaaagctgtg aagatacggg agagaactcc agaagacatt 300  
32 tttaaacctt ctaatgggat cattcatcat tttaaaacca tgcaccgata cacactggaa 360  
33 atgttcagaa cttgccagtt ttgtcctcag ttccgggaga tcatccacaa agccctcatc 420  
34 gacagaaaca tccaggccac cctggaaagc cagaagaaac tcaactgggtg tcgagaagtc 480  
35 cggaagcttg tggcgctgaa aacgaacggt gacggcaatt gcctcatgca tgccacttct 540  
36 cagtacatgt ggggcgttca ggacacagac ttgggtactga ggaaggcgct gttcagcacg 600  
37 ctcaaggaaa cagacacacg caactttaaa ttccgctggc aactggagtc tctcaaattc 660  
38 caggaatttg ttgaaacggg gctttgctat gatactcgga actggaatga tgaatgggac 720  
39 aatcttatca aaatggcttc cacagacaca cccatggccc gaagtggact tcagtacaac 780  
40 tcaactggaag aaatacacat atttgcctt tgcaacatcc tcagaaggcc aatcattgtc 840  
41 atttcagaca aaatgctaag aagtttgga tcaggttcca atttcgcccc tttgaaagt 900  
42 ggtggaattt acttgccctc ccaactggcct gccagggaat gctacagata cccattgtt 960  
43 ctccgctatg acagccatca tttgttacc ttggtgacct tgaaggacag tgggcctgaa 1020  
44 atccgagctg ttccacttgt taacagagac cggggaagat ttgaagactt aaaagttcac 1080  
45 tttttgacag atcctgaaaa tgagatgaag gagaagctct taaaagagta cttaatggtg 1140  
46 atagaaatcc ccgtccaagg ctgggacctt ggcacaactc atctcatcaa tgccgcaaag 1200  
47 ttggatgaag ctaacttacc aaaagaaatc aatctggtag atgattactt tgaacttggt 1260  
48 cagcatgagt acaagaaatg gcaggaaaac agcgagcagg ggaggagaga ggggcacgcc 1320  
49 cagaatccca tggaaccttc cgtgccccag ctttctctca tggatgtaaa atgtgaaacg 1380  
50 cccaactgcc ccttcttcat gtctgtgaac acccagcctt tatgccatga gtgctcagag 1440  
51 aggcggcaaa agaatacaaaa caaactccca aagctgaact ccaagccggg ccctgagggg 1500  
52 ctccctggca tggcgctcgg ggctctcgg ggagaagcct atgagccctt ggcgtggaac 1560  
53 cctgaggagt ccaactgggg gcctcattcg gccccaccga cagcaccag cccttttctg 1620  
54 ttcagtgaga ccaactgcat gaagtgcagg agccccggct gcccttcac actgaatgtg 1680  
55 cagcacaacg gatattgtga acgttgccac aacgcccggc aacttcacgc cagccacgcc 1740  
56 ccagaccaca caaggcactt ggatcccggg aagtgccaa cctgcctcca ggatgttacc 1800  
57 aggacattta atgggatctg cagtacttgc ttcaaaagga ctacagcaga ggcctcctcc 1860

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001

TIME: 08:56:25

Input Set : A:\Seqlist.txt

Output Set : N:\CRF3\12132001\I003608.raw

```

58 agcctcagca ccagcctccc tccttcctgt caccagcggtt ccaagtcaga tccctcgcg 1920
59 ctcgctccgga gcccctcccc gcattcttgc cacagagctg gaaacgacgc ccctgctggc 1980
60 tgctgtctc aagctgcacg gactcctggg gacaggacgg ggacgagcaa gtgcagaaaa 2040
61 gccggctgcy tgtatttttg gactccagaa aacaaggggt tttgcacact gtgtttcatc 2100
62 gagtacagag aaaacaaaca ttttgcctgt gcctcaggga aagtcagtcc cacagcgctc 2160
63 aggttccaga acaccattcc gtgcctgggg agggaatgcy gcacccttgg aagcaccatg 2220
64 tttgaaggat actgccagaa gtgtttcatt gaagtcaga atcagagatt tcatgaggcc 2280
65 aaaaggacag aagagcaact gagatcgagc cagcgacagag atgtgcctcg aaccacacaa 2340
66 agcacctcaa ggcccaagtg cggccggggc tcctgcaaga acatcctggc ctgcccagc 2400
67 gaggagctct gcatggagtg tcagcatccc aaccagagga tgggcccctg ggcccaccgg 2460
68 ggtgagcctg cccccaaga ccccccaag cagcgttgcc gggcccccg cgtgtatcat 2520
69 tttggcaatg ccaagtcaa cggctactgc aacgaatgct ttcagttcaa gcagatgtat 2580
70 ggctaaccgg aaacaggtgg gtcacctcct gcaagaagtg gggcctcgag ctgtcagtca 2640
71 tcatggtgct atcctctgaa cccctcagct gccactgcaa cagtgggctt aagggtgtct 2700
72 gagcaggaga ggaaagataa gctcttcgtg gtgccacga tgcctcaggt tggttaaccg 2760
73 ggagtgttcc caggtggcct tagaaagcaa agcttgtaac tggcaaggga tgatgtcaga 2820
74 ttcagcccaa ggttctcct ctctaccaa gcaggaggcc aggaacttct ttggacttgg 2880
75 aaggtgtgcy gggactggcc gagccccctg caccctgcgc atcaggactg cttcatcgtc 2940
76 ttggctgaga aagggaaaag acacacaagt cgcgtgggtt ggagaagcca gagccattcc 3000
77 acctcccctc cccagcatc tctcagagat gtgaagccag atcctcatgg cagcgaggcc 3060
78 ctctgcaaga agctcaagga agctcaggga aaatggacgt attcagagag tgtttgtagt 3120
79 tcatggtttt tccctacctg cccggttctt ttcctgagga cccggcagaa atgcagaacc 3180
80 atccatggac tgtgattctg aggtgctga gactgaacat gttcacattg acagaaaaac 3240
81 aagctgctct ttataatatg caccttttaa aaaattagaa tattttactg ggaagacgtg 3300
82 taactctttg ggttattact gtctttactt ctaaagaagt tagcttgaac tgaggagtaa 3360
83 aagtgtgtac atatataata tacccttaca ttatgtatga gggatttttt taaattatat 3420
84 tgaaatgctg cctagaagt acaataggaa ggctaaataa taataacctg tttctggtt 3480
85 gttgttgggg catgagcttg tgtatacact gcttgcataa actcaaccag ctgccttttt 3540
86 aaaggagct ctagtctttt ttgtgtaatt cactttattt attttattac aaacttcaag 3600
87 attatttaag tgaagatatt tcttcagctc tggggaaaat gccacagtgt tctcctgaga 3660
88 gaacatcctt gctttgagtc aggtgtggg caagtctctg accacaggga gtaaattggc 3720
89 ctctttgata cacttttgct tgcctcccca ggaaagaagg aattgcatcc aaggtataca 3780
90 tacatattca tcgatgtttc gtgcttctcc ttatgaaact ccagctatgt aataaaaaac 3840
91 tatactctgt gttctgttaa tgcctctgag tgcctacct ccttgagat gagataggga 3900
92 aggagcagg atgagactgg caatggtcac agggaaagat gtggcctttt gtgatgggtt 3960
93 tattttctgt taacactgtg tcctgggggg gctgggaagt cccctgcac ccatgggtacc 4020
94 ctggtattgg gacagcaaaa gccagtaacc atgagtatga ggaaatctct ttctgttgc 4080
95 ggcttacagt ttctctgtgt gctttgtggt tgcgtgcata tttgctctag aagaaaaaaa 4140
96 aaaaaaggag gggaaatgca ttttccccag agataaaggc tgccattttg ggggtctgta 4200
97 cttatggcct gaaaatattt gtgatccata actctacaca gcctttactc atactattag 4260
98 gcacacttcc ccttagagc cccctaagtt tttccagac gaatctttat aatttctttc 4320
99 caaagatacc aaataaactt cagtgttttc atctaattct cttaaagttg atatcttaat 4380
100 attttgtgtt gatcattatt tccattctta atgtgaaaaa aagtaattat ttatacttat 4440
101 tataaaaagt atttgaaatt tgcacattta attgtcccta atagaaagcc acctattctt 4500
102 tgttggtatt ctcaagttt ttctaaataa atgtaacttt tcacaagagt caacattaaa 4560
103 aaataaatta tttaaaaaaa aaaaaaaa 4588
105 <210> SEQ ID NO: 2
106 <211> LENGTH: 790
107 <212> TYPE: PRT

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001  
TIME: 08:56:25

Input Set : A:\Seqlist.txt  
Output Set: N:\CRF3\12132001\I003608.raw

```

108 <213> ORGANISM: Human
110 <400> SEQUENCE: 2
111 Met Ala Glu Gln Val Leu Pro Gln Ala Leu Tyr Leu Ser Asn Met Arg
112 1 5 10 15
113 Lys Ala Val Lys Ile Arg Glu Arg Thr Pro Glu Asp Ile Phe Lys Pro
114 20 25 30
115 Thr Asn Gly Ile Ile His His Phe Lys Thr Met His Arg Tyr Thr Leu
116 35 40 45
117 Glu Met Phe Arg Thr Cys Gln Phe Cys Pro Gln Phe Arg Glu Ile Ile
118 50 55 60
119 His Lys Ala Leu Ile Asp Arg Asn Ile Gln Ala Thr Leu Glu Ser Gln
120 65 70 75 80
121 Lys Lys Leu Asn Trp Cys Arg Glu Val Arg Lys Leu Val Ala Leu Lys
122 85 90 95
123 Thr Asn Gly Asp Gly Asn Cys Leu Met His Ala Thr Ser Gln Tyr Met
124 100 105 110
125 Trp Gly Val Gln Asp Thr Asp Leu Val Leu Arg Lys Ala Leu Phe Ser
126 115 120 125
127 Thr Leu Lys Glu Thr Asp Thr Arg Asn Phe Lys Phe Arg Trp Gln Leu
128 130 135 140
129 Glu Ser Leu Lys Ser Gln Glu Phe Val Glu Thr Gly Leu Cys Tyr Asp
130 145 150 155 160
131 Thr Arg Asn Trp Asn Asp Glu Trp Asp Asn Leu Ile Lys Met Ala Ser
132 165 170 175
133 Thr Asp Thr Pro Met Ala Arg Ser Gly Leu Gln Tyr Asn Ser Leu Glu
134 180 185 190
135 Glu Ile His Ile Phe Val Leu Cys Asn Ile Leu Arg Arg Pro Ile Ile
136 195 200 205
137 Val Ile Ser Asp Lys Met Leu Arg Ser Leu Glu Ser Gly Ser Asn Phe
138 210 215 220
139 Ala Pro Leu Lys Val Gly Gly Ile Tyr Leu Pro Leu His Trp Pro Ala
140 225 230 235 240
141 Gln Glu Cys Tyr Arg Tyr Pro Ile Val Leu Gly Tyr Asp Ser His His
142 245 250 255
143 Phe Val Pro Leu Val Thr Leu Lys Asp Ser Gly Pro Glu Ile Arg Ala
144 260 265 270
145 Val Pro Leu Val Asn Arg Asp Arg Gly Arg Phe Glu Asp Leu Lys Val
146 275 280 285
147 His Phe Leu Thr Asp Pro Glu Asn Glu Met Lys Glu Lys Leu Leu Lys
148 290 295 300
149 Glu Tyr Leu Met Val Ile Glu Ile Pro Val Gln Gly Trp Asp His Gly
150 305 310 315 320
151 Thr Thr His Leu Ile Asn Ala Ala Lys Leu Asp Glu Ala Asn Leu Pro
152 325 330 335
153 Lys Glu Ile Asn Leu Val Asp Asp Tyr Phe Glu Leu Val Gln His Glu
154 340 345 350
155 Tyr Lys Lys Trp Gln Glu Asn Ser Glu Gln Gly Arg Arg Glu Gly His
156 355 360 365
157 Ala Gln Asn Pro Met Glu Pro Ser Val Pro Gln Leu Ser Leu Met Asp

```

12/13/01

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001

TIME: 08:56:25

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12132001\I003608.raw

```

158      370      375      380
159 Val Lys Cys Glu Thr Pro Asn Cys Pro Phe Phe Met Ser Val Asn Thr
160 385      390      395      400
161 Gln Pro Leu Cys His Glu Cys Ser Glu Arg Arg Gln Lys Asn Gln Asn
162      405      410      415
163 Lys Leu Pro Lys Leu Asn Ser Lys Pro Gly Pro Glu Gly Leu Pro Gly
164      420      425      430
165 Met Ala Leu Gly Ala Ser Arg Gly Glu Ala Tyr Glu Pro Leu Ala Trp
166      435      440      445
167 Asn Pro Glu Glu Ser Thr Gly Gly Pro His Ser Ala Pro Pro Thr Ala
168      450      455      460
169 Pro Ser Pro Phe Leu Phe Ser Glu Thr Thr Ala Met Lys Cys Arg Ser
170 465      470      475      480
171 Pro Gly Cys Pro Phe Thr Leu Asn Val Gln His Asn Gly Phe Cys Glu
172      485      490      495
173 Arg Cys His Asn Ala Arg Gln Leu His Ala Ser His Ala Pro Asp His
174      500      505      510
175 Thr Arg His Leu Asp Pro Gly Lys Cys Gln Ala Cys Leu Gln Asp Val
176      515      520      525
177 Thr Arg Thr Phe Asn Gly Ile Cys Ser Thr Cys Phe Lys Arg Thr Thr
178      530      535      540
179 Ala Glu Ala Ser Ser Ser Leu Ser Thr Ser Leu Pro Pro Ser Cys His
180 545      550      555      560
181 Gln Arg Ser Lys Ser Asp Pro Ser Arg Leu Val Arg Ser Pro Ser Pro
182      565      570      575
183 His Ser Cys His Arg Ala Gly Asn Asp Ala Pro Ala Gly Cys Leu Ser
184      580      585      590
185 Gln Ala Ala Arg Thr Pro Gly Asp Arg Thr Gly Thr Ser Lys Cys Arg
186      595      600      605
187 Lys Ala Gly Cys Val Tyr Phe Gly Thr Pro Glu Asn Lys Gly Phe Cys
188      610      615      620
189 Thr Leu Cys Phe Ile Glu Tyr Arg Glu Asn Lys His Phe Ala Ala Ala
190 625      630      635      640
191 Ser Gly Lys Val Ser Pro Thr Ala Ser Arg Phe Gln Asn Thr Ile Pro
192      645      650      655
193 Cys Leu Gly Arg Glu Cys Gly Thr Leu Gly Ser Thr Met Phe Glu Gly
194      660      665      670
195 Tyr Cys Gln Lys Cys Phe Ile Glu Ala Gln Asn Gln Arg Phe His Glu
196      675      680      685
197 Ala Lys Arg Thr Glu Glu Gln Leu Arg Ser Ser Gln Arg Arg Asp Val
198      690      695      700
199 Pro Arg Thr Thr Gln Ser Thr Ser Arg Pro Lys Cys Ala Arg Ala Ser
200 705      710      715      720
201 Cys Lys Asn Ile Leu Ala Cys Arg Ser Glu Glu Leu Cys Met Glu Cys
202      725      730      735
203 Gln His Pro Asn Gln Arg Met Gly Pro Gly Ala His Arg Gly Glu Pro
204      740      745      750
205 Ala Pro Glu Asp Pro Pro Lys Gln Arg Cys Arg Ala Pro Ala Cys Asp
206      755      760      765

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001

TIME: 08:56:25

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12132001\I003608.raw

207 His Phe Gly Asn Ala Lys Cys Asn Gly Tyr Cys Asn Glu Cys Phe Gln  
 208 770 775 780

209 Phe Lys Gln Met Tyr Gly  
 210 785 790

213 &lt;210&gt; SEQ ID NO: 3

214 &lt;211&gt; LENGTH: 1224

215 &lt;212&gt; TYPE: DNA

216 &lt;213&gt; ORGANISM: Human

218 &lt;220&gt; FEATURE:

219 &lt;221&gt; NAME/KEY: misc\_feature

220 &lt;222&gt; LOCATION: 36, 91, 645, 655, 660, 671, 672

221 &lt;223&gt; OTHER INFORMATION: n = A,T,C or G

223 &lt;221&gt; NAME/KEY: misc\_feature

224 &lt;222&gt; LOCATION: 36, 91, 645, 655, 660, 671, 672

225 &lt;223&gt; OTHER INFORMATION: n = A,T,C or G

227 &lt;221&gt; NAME/KEY: allele

228 &lt;222&gt; LOCATION: (0)...(0)

230 &lt;400&gt; SEQUENCE: 3

W--> 231 tctggatcga tctggagctc cgggaatttc cctggncggg gactccgggc tttccagccc 60  
 W--> 232 caaccatgca taaaaggggt tgcgcgttct nggagagcca cagagccgg gccacaggca 120  
 233 gtctcttgcc agctcttct ctcctctcac agccgccaga cccgcctgct gagcccccat 180  
 234 ggcccgcgct gctctctccg ccgccccag caatccccgg ctctcgagag tggcgctgct 240  
 235 gtctctgctc ctggtagccg ctggccggcg cgcagcagga gcgcccctgg ccaactgaact 300  
 236 gcgctgccag tgcttgca cctgcaggg aattcacctc aagaacatcc aaagtgtgaa 360  
 237 ggtgaagtcc cccggacccc actgcgcca aaccgaagtc atagccacac tcaagaatgg 420  
 238 gcagaaagct tgtctcaacc ccgcacgcc catggttaag aaaatcatcg aaaagatgct 480  
 239 gaaaaatggc aaatccaact gaccagaagg aaggagggaag cttattgggtg gctgttcctg 540  
 240 aaggaggccc tgcccttaca ggaacagaag aggaaagaga gacacagctg cagaggccac 600  
 W--> 241 ctgggattgc gcctaattgtg tttgagcatc acttaggaga aggcncggat taatnaattn 660  
 W--> 242 attaatttat nnattggttg gttttagaag attctatggt aatattttat gtgtaaaata 720  
 243 aggttatgat tgaatctact tgcacactct cccattatat ttattgttta ttttaggtca 780  
 244 aacccaagtt agttcaatcc tgattcatat ttaatttgaa gatagaaggt ttgcagatat 840  
 245 tctctagtca tttgttaata tttcttcgtg atgacatatc acatgtcagc cactgtgata 900  
 246 gaggtgagg aatccaagaa aatggccagt aagatcaatg tgacggcagg gaaatgtatg 960  
 247 tgtgtctatt ttgtaactgt aaagatgaat gtcagttggt atttattgaa atgatttcac 1020  
 248 agtgtgtggt caacatttct catgttgaag ctttaagaac taaaatgttc taaatatccc 1080  
 249 ttggacattt tatgtcttct ttgtaaggca tactgccttg tttaatgtta attatgcagt 1140  
 250 gtttccctct gtgtagagc agagagggtt cgaatattat tgatgttttc acaaagaaca 1200  
 251 ggaaaataaa atatttaaaa atat 1224

*delete - duplicate*

253 &lt;210&gt; SEQ ID NO: 4

254 &lt;211&gt; LENGTH: 107

255 &lt;212&gt; TYPE: PRT

256 &lt;213&gt; ORGANISM: Human

258 &lt;400&gt; SEQUENCE: 4

259 Met Ala Arg Ala Ala Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu

260 1 5 10 15

261 Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Gly Arg Arg Ala

262 20 25 30

263 Ala Gly Ala Pro Leu Ala Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/003,608

DATE: 12/13/2001

TIME: 08:56:26

Input Set : A:\Seqlist.txt

Output Set: N:\CRF3\12132001\I003608.raw

L:7 M:201 W: Mandatory field data missing, TITLE INVENTION  
L:14 M:270 C: Current Application Number differs, Replaced Current Application No  
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:232 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:241 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:486 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11